

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/757,356
Source: IFWO
Date Processed by STIC: 10-18-04

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/757,356

TIME: 10:04:57

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\10182004\J757356.raw

(pg. 6)

4 <110> APPLICANT: Labas, Yulii Aleksandrovich
 5 Gurskaya, Nadezda Georgievna
 6 Yanushevich, Yuriy
 7 Fradkov, Arcady Fedorovich
 8 Lukyanov, Konstantin
 9 Lukyanov, Sergey
 10 Matz, Mikhail Vladimirovich
 12 <120> TITLE OF INVENTION: NOVEL CHROMOPHORES/FLUOROPHORES AND
 13 METHODS FOR USING THE SAME
 16 <130> FILE REFERENCE: CLON-090
 18 <140> CURRENT APPLICATION NUMBER: 10/757,356
 19 <141> CURRENT FILING DATE: 2004-01-13
 21 <150> PRIOR APPLICATION NUMBER: 60/332,980
 22 <151> PRIOR FILING DATE: 2001-11-13
 24 <150> PRIOR APPLICATION NUMBER: PCT/US02/36499
 25 <151> PRIOR FILING DATE: 2002-11-12
 27 <160> NUMBER OF SEQ ID NOS: 28
 29 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 31 <210> SEQ ID NO: 1
 32 <211> LENGTH: 868
 33 <212> TYPE: DNA
 34 <213> ORGANISM: Heteractis crispa
 36 <400> SEQUENCE: 1
 37 attttgga ggtgttcaac caagcaaatt taagaagtca tcatctttat ctcaagtcagg 60
 38 aaaatgtgtt cttacatcaa agaaaccatg caaagtaagg ttacatgga aggaaaagtt 120
 39 aacgaccaca acttcaagt cactgcagaa ggaaaaggag aaccatacaa aggctcacia 180
 40 agcctgacga tcaccgtaac tgaaggaggt cctctgccat ttgccttcga cattctttca 240
 41 cacgcctttc gatattggca taagggtgtt gccaaagtacc ccaaagatca tcctgatttt 300
 42 ttttaagcagt ctcttcctga aggtttttact tgggaaagag taagcaacta tgaggacgga 360
 43 ggagtcctta ccgttaaaaca agaaactagt ctggaggagg attgcattat ttgcaaaatt 420
 44 aaagcacatg gcaactaactt cccgcgagat ggtccggtga tgcaaaaacg gaccaatgga 480
 45 tgggagccat caactgaaac gggtattcca cgggggtggag gaattctgat gcgcgatgtg 540
 46 cccgcactga agctgcttgg taacaaagga catcttctct gcgtcatgga aacaacttac 600
 47 aagtcaaaaa aaaaagggtga acctgccaaa ccgcactttc atcatttgag aatggagaag 660
 48 gatagtgtta gtgacgatga gaagaccatt gagcagcacg agaatgtgag ggcaagctac 720
 49 ttcaatgata gtggaaaatg atcatttcct tattgatttc aatgttaggg cattcagttt 780
 50 ccaaattttc ttagacacag tcttttcctt tagcttcgta gctacttac ccatgttttg 840
 51 ttgaagtcaa taaatagcta agcactac 868
 53 <210> SEQ ID NO: 2
 54 <211> LENGTH: 225
 55 <212> TYPE: PRT
 56 <213> ORGANISM: Heteractis crispa
 58 <400> SEQUENCE: 2

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/757,356

DATE: 10/18/2004

TIME: 10:04:57

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\10182004\J757356.raw

```

59 Met Cys Ser Tyr Ile Lys Glu Thr Met Gln Ser Lys Val Tyr Met Glu
60 1 5 10 15
61 Gly Lys Val Asn Asp His Asn Phe Lys Cys Thr Ala Glu Gly Lys Gly
62 20 25 30
63 Glu Pro Tyr Lys Gly Ser Gln Ser Leu Thr Ile Thr Val Thr Glu Gly
64 35 40 45
65 Gly Pro Leu Pro Phe Ala Phe Asp Ile Leu Ser His Ala Phe Arg Tyr
66 50 55 60
67 Gly Asn Lys Val Phe Ala Lys Tyr Pro Lys Asp His Pro Asp Phe Phe
68 65 70 75 80
69 Lys Gln Ser Leu Pro Glu Gly Phe Thr Trp Glu Arg Val Ser Asn Tyr
70 85 90 95
71 Glu Asp Gly Gly Val Leu Thr Val Lys Gln Glu Thr Ser Leu Glu Gly
72 100 105 110
73 Asp Cys Ile Ile Cys Lys Ile Lys Ala His Gly Thr Asn Phe Pro Ala
74 115 120 125
75 Asp Gly Pro Val Met Gln Lys Arg Thr Asn Gly Trp Glu Pro Ser Thr
76 130 135 140
77 Glu Thr Val Ile Pro Arg Gly Gly Gly Ile Leu Met Arg Asp Val Pro
78 145 150 155 160
79 Ala Leu Lys Leu Leu Gly Asn Lys Gly His Leu Leu Cys Val Met Glu
80 165 170 175
81 Thr Thr Tyr Lys Ser Lys Lys Lys Gly Glu Pro Ala Lys Pro His Phe
82 180 185 190
83 His His Leu Arg Met Glu Lys Asp Ser Val Ser Asp Asp Glu Lys Thr
84 195 200 205
85 Ile Glu Gln His Glu Asn Val Arg Ala Ser Tyr Phe Asn Asp Ser Gly
86 210 215 220
87 Lys
88 225
91 <210> SEQ ID NO: 3
92 <211> LENGTH: 845
93 <212> TYPE: DNA
94 <213> ORGANISM: Dendronephthya sp
96 <400> SEQUENCE: 3
97 catatcgaga aagttgtgaa accaaattct tactctactt ttactaccat gaatctgatt 60
98 aaagaagata tgaggggttaa ggtgcatatg gaaggggaatg taaacgggca tgcttttgtg 120
99 attgaagggg aaggaaaagg aaggccctac gaagggaacac agaccttgaa cctgacagtg 180
100 aaagaaggcg cgctctctccc attttcttac gacatcttga caacagcatt gcactacgga 240
101 aacagagtat tcaactgaata cccagcagat atcacggatt atttcaagca atcatttctt 300
102 gaaggatatt cctgggaaag aaccatgact tatgaagaca agggcatttg taccatcaga 360
103 agcgacataa gcttggaagg tgactgcttt ttccaaaaca ttcgttttaa tgggatgaac 420
104 tttcccccaa atggtccagt tatgcagaag aaaactttga agtggaacc atccacagag 480
105 aagctgcacg tgcgtgatgg gttgcttgtc ggtaatatta acatggctct gctgcttgaa 540
106 ggaggtggac attacctgtg tgacttcaaa actacttaca aagcgaagaa ggttggttcag 600
107 ttgccagatt atcattttgt ggaccatcgc attgagatct tgagtaatga cagcgattac 660
108 aacaaagtga agctgtacga gcatgggggt gctcgctatt ctccgttgcc caagtcaggc 720
109 ctggtagagg ttcaaggga agccataatg actgcataga taaacatgta gtgaagacca 780
110 catactcggg attagagttt agggattggg agttgtggta gattctagcc tacaaatttt 840

```

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/757,356

TIME: 10:04:57

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\10182004\J757356.raw

```

111 ttggg
113 <210> SEQ ID NO: 4
114 <211> LENGTH: 236
115 <212> TYPE: PRT
116 <213> ORGANISM: Dendronephthya sp
118 <400> SEQUENCE: 4
119 Met Asn Leu Ile Lys Glu Asp Met Arg Val Lys Val His Met Glu Gly
120 1 5 10 15
121 Asn Val Asn Gly His Ala Phe Val Ile Glu Gly Glu Gly Lys Gly Arg
122 20 25 30
123 Pro Tyr Glu Gly Thr Gln Thr Leu Asn Leu Thr Val Lys Glu Gly Ala
124 35 40 45
125 Pro Leu Pro Phe Ser Tyr Asp Ile Leu Thr Thr Ala Leu His Tyr Gly
126 50 55 60
127 Asn Arg Val Phe Thr Glu Tyr Pro Ala Asp Ile Thr Asp Tyr Phe Lys
128 65 70 75 80
129 Gln Ser Phe Pro Glu Gly Tyr Ser Trp Glu Arg Thr Met Thr Tyr Glu
130 85 90 95
131 Asp Lys Gly Ile Cys Thr Ile Arg Ser Asp Ile Ser Leu Glu Gly Asp
132 100 105 110
133 Cys Phe Phe Gln Asn Ile Arg Phe Asn Gly Met Asn Phe Pro Pro Asn
134 115 120 125
135 Gly Pro Val Met Gln Lys Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu
136 130 135 140
137 Lys Leu His Val Arg Asp Gly Leu Leu Val Gly Asn Ile Asn Met Ala
138 145 150 155 160
139 Leu Leu Leu Glu Gly Gly His Tyr Leu Cys Asp Phe Lys Thr Thr
140 165 170 175
141 Tyr Lys Ala Lys Lys Val Val Gln Leu Pro Asp Tyr His Phe Val Asp
142 180 185 190
143 His Arg Ile Glu Ile Leu Ser Asn Asp Ser Asp Tyr Asn Lys Val Lys
144 195 200 205
145 Leu Tyr Glu His Gly Val Ala Arg Tyr Ser Pro Leu Pro Lys Ser Gly
146 210 215 220
147 Leu Val Glu Val Gln Gly Lys Ala Ile Met Thr Ala
148 225 230 235
151 <210> SEQ ID NO: 5
152 <211> LENGTH: 851
153 <212> TYPE: DNA
154 <213> ORGANISM: Zoanthus sp
156 <400> SEQUENCE: 5
157 gagttgagtt ctcgacttca gttgtatcac ttttgacgta tcaagtgatc tattctcaac 60
158 atggccatt caaagcacgg actaacagat gacatgacaa tgcatttccg tatggaaggg 120
159 tgcgtcgatg gacataagtt tgtaatcgag ggcaacggca atggaaatcc tttcaaaggg 180
160 aaacagttta ttaatctgtg tgtgattgaa ggaggaccac tgccattctc cgaagacata 240
161 ttgtctgctg cgtttgacta cggaaacagg ctcttcaactg aatatcctga aggcatagtt 300
162 gactatttca agaactcgtg tcctgctgga tatacgtggc acaggtcttt tcgctttgaa 360
163 gatggagcag tttgcatatg cagtgcagat ataacagtaa atgttaggga aaactgcatt 420
164 tatcatgagt ccacgtttta tggagtgaac tttcctgctg atggacctgt gatgaaaaag 480

```

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/757,356

TIME: 10:04:57

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\10182004\J757356.raw

```

165 atgacaacta attgggaacc gtcctgcgag aaaatcatat caataaatag tcagaagata 540
166 ttaaaagggg atgtctccat gtacctcctt ctgaaggatg gtgggcgtta ccgctgccag 600
167 tttgacacaa ttacaaaagc aaagactgag ccaaaagaaa tgccggactg gcacttcac 660
168 cagcataagc tcaaccgtga agaccgcagc gatgctaaga atcagaaatg gcaactgata 720
169 gaacatgcta ttgcatcccg atctgcttta ccctgataac aaaggagttg ctattgcatg 780
170 tgcattgcta ttacgtgat aaaaatgtag ttttaacatg caattgtatg tgcattgaca 840
171 ttacctgat a 851

```

173 <210> SEQ ID NO: 6

174 <211> LENGTH: 231

175 <212> TYPE: PRT

176 <213> ORGANISM: Zoanthus sp

178 <400> SEQUENCE: 6

```

179 Met Ala His Ser Lys His Gly Leu Thr Asp Asp Met Thr Met His Phe
180 1 5 10 15
181 Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Glu Gly Asn
182 20 25 30
183 Gly Asn Gly Asn Pro Phe Lys Gly Lys Gln Phe Ile Asn Leu Cys Val
184 35 40 45
185 Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Ala
186 50 55 60
187 Phe Asp Tyr Gly Asn Arg Leu Phe Thr Glu Tyr Pro Glu Gly Ile Val
188 65 70 75 80
189 Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp His Arg Ser
190 85 90 95
191 Phe Arg Phe Glu Asp Gly Ala Val Cys Ile Cys Ser Ala Asp Ile Thr
192 100 105 110
193 Val Asn Val Arg Glu Asn Cys Ile Tyr His Glu Ser Thr Phe Tyr Gly
194 115 120 125
195 Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn
196 130 135 140
197 Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Ile Asn Ser Gln Lys Ile
198 145 150 155 160
199 Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
200 165 170 175
201 Tyr Arg Cys Gln Phe Asp Thr Ile Tyr Lys Ala Lys Thr Glu Pro Lys
202 180 185 190
203 Glu Met Pro Asp Trp His Phe Ile Gln His Lys Leu Asn Arg Glu Asp
204 195 200 205
205 Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Ile Glu His Ala Ile
206 210 215 220
207 Ala Ser Arg Ser Ala Leu Pro
208 225 230

```

211 <210> SEQ ID NO: 7

212 <211> LENGTH: 1178

213 <212> TYPE: DNA

214 <213> ORGANISM: Scolymia cubensis

216 <400> SEQUENCE: 7

```

217 tgtgacattc agtcatatag gagcctctat cggagctgag gtccattca ccgttgtgat 60
218 ttggacggga gcagatcgag aacaacmagg gctgtacgag tctgataatt tactttacat 120

```

RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/757,356

TIME: 10:04:57

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\10182004\J757356.raw

```

219 ctaccaacat gcagcgtgct gggatgaagg ttaaggaaca tatgaagatc aaactgcgta 180
220 tgggaggtac tgtaaacgga aagcatttcg cggttaatgg gacaggagac ggctaccctt 240
221 atcagggaaa acagattttg aaacttatcg tcgaaggcag cgaacctctg cctttcgctt 300
222 ttgatattctt gtcagcagca ttccagtatg gcaacagggc attcaccgaa tacccaacag 360
223 agatagcaga ctattttcaag cagtcgtttg agtttgccga ggggttctcc tgggaacgaa 420
224 gtttcacttt cgaagatggg gccatttgcg tcgccaccaaa cgatataacg atggttggtg 480
225 gtgagtttca gtatgatatt cgatttgatg gtctgaactt ccctgaagat ggtccagtga 540
226 tgcaaaagaa aaccgtaaaa tgggagccat ccactgagat aatgtatatg caaaatggag 600
227 tgctgaaggg tgaggttaac atggctctgt tgcttcaaga caaaagccat taccgttgcg 660
228 acctcaaaac tacttacaaa gctaagaata atgtgccgca tcctccaggc taccactatg 720
229 tggatcactg cattgaaata ctggaagaac gtaaggatca cgtaaagctg cgggagcatg 780
230 ctaaaagctcg ttctagcctg tcacctacca gtgcaaaaga acgaaaggct taggtgatag 840
231 tcaaaaagac aacaagacga aaatgaaggg tgttcattgt tagaatttga tattttcgat 900
232 tcaatgattc gttaagggat ttgctagagg ctgactaaca ggtaaacatc ataaggatag 960
233 agattycgtt gcggagttag aaccttwata ttttccgaat tccamctaga gtcggttgaga 1020
234 aattttattag agactagctt tagagttact tttgtggaaa aaaagggttc cattttttgc 1080
235 gttattacag catttaaagc ataggaatag agattcgggt atggaaaata acagtaggaa 1140
236 aatacgttgt gaaaataaac ttgttgctga aaaaaaaa 1178

```

238 <210> SEQ ID NO: 8

239 <211> LENGTH: 234

240 <212> TYPE: PRT

241 <213> ORGANISM: Scolymia cubensis

243 <400> SEQUENCE: 8

```

244 Met Gln Arg Ala Gly Met Lys Val Lys Glu His Met Lys Ile Lys Leu
245 1 5 10 15
246 Arg Met Gly Gly Thr Val Asn Gly Lys His Phe Ala Val Asn Gly Thr
247 20 25 30
248 Gly Asp Gly Tyr Pro Tyr Gln Gly Lys Gln Ile Leu Lys Leu Ile Val
249 35 40 45
250 Glu Gly Ser Glu Pro Leu Pro Phe Ala Phe Asp Ile Leu Ser Ala Ala
251 50 55 60
252 Phe Gln Tyr Gly Asn Arg Ala Phe Thr Glu Tyr Pro Thr Glu Ile Ala
253 65 70 75 80
254 Asp Tyr Phe Lys Gln Ser Phe Glu Phe Gly Glu Gly Phe Ser Trp Glu
255 85 90 95
256 Arg Ser Phe Thr Phe Glu Asp Gly Ala Ile Cys Val Ala Thr Asn Asp
257 100 105 110
258 Ile Thr Met Val Gly Gly Glu Phe Gln Tyr Asp Ile Arg Phe Asp Gly
259 115 120 125
260 Leu Asn Phe Pro Glu Asp Gly Pro Val Met Gln Lys Lys Thr Val Lys
261 130 135 140
262 Trp Glu Pro Ser Thr Glu Ile Met Tyr Met Gln Asn Gly Val Leu Lys
263 145 150 155 160
264 Gly Glu Val Asn Met Ala Leu Leu Leu Gln Asp Lys Ser His Tyr Arg
265 165 170 175
266 Cys Asp Leu Lys Thr Thr Tyr Lys Ala Lys Asn Asn Val Pro His Pro
267 180 185 190
268 Pro Gly Tyr His Tyr Val Asp His Cys Ile Glu Ile Leu Glu Glu Arg
269 195 200 205

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/757,356

DATE: 10/18/2004
TIME: 10:04:58

Input Set : D:\Seqlist.txt
Output Set: N:\CRF4\10182004\J757356.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:28; Xaa Pos. 65,144

VERIFICATION SUMMARY

DATE: 10/18/2004

PATENT APPLICATION: US/10/757,356

TIME: 10:04:58

Input Set : D:\Seqlist.txt

Output Set: N:\CRF4\10182004\J757356.raw

L:871 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:64
M:341 Repeated in SeqNo=28